

Python для структур и не только

Задание 1. ProDy и B-факторы часть 1

Мною был взят белок со структурой 5RYR¹, использованный ранее в [практикуме 2](#).

Сначала я использовала ProDy для нахождения остатка, средний B-фактор атомов которого был максимален. Остаток с наибольшим средним B-фактором (91.593) – Cys-245 цепи А. Его изображение в модели структуры 5RYR можно увидеть на Рис. 1. Данный остаток расположен на петле, его радикал «смотрит» в раствор. Значения B-фактора основных атомов находятся в интервале $\sim (91 - 94)$, самым подвижным оказался атом кислорода: его B-фактор - 93.69. У крайних атомов радикала SG и CB значения B-факторов – 91.2 и 91.1, соответственно.

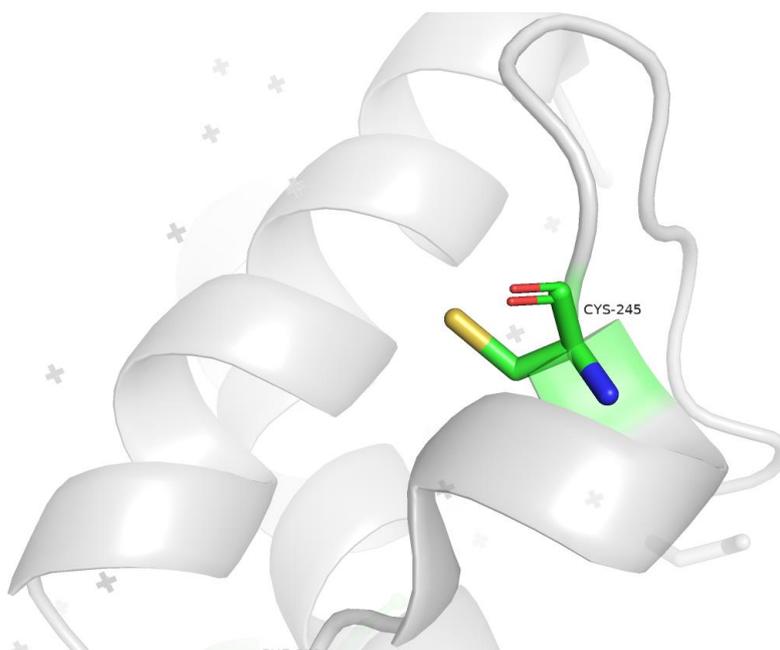


Рис. 1. Остаток Cys-245 (с наибольшим B-фактором)

Далее аналогичным способом я нашла остаток с наименьшим средним B-фактором (20.806). Им оказался Phe-286 цепи А (см. Рис. 2). Данный остаток расположен на α -спирали внутри глобулы белка, его боковой радикал направлен внутрь ансамбля из трех α -спиралей. Остаток относительно неподвижен, значения B-факторов всех его атомов находятся в пределах $\sim(19-22)$.

¹ Кристаллографическую расшифровку [5RYR](#) имеет белок EPB41L3. Он локализуется в цитоплазме, в цитоскелете и в плазматической мембране [1]. EPB41L3 (erythrocyte membrane protein band 4.1 like 3) – мембранный белок эритроцита. Это опухолевый супрессор, который подавляет пролиферацию клетки и способствует апоптозу. Данный белок регулирует активность протеин аргинин N-метилтрансферазы (включая PRMT3 и PRMT5) [1]. А PRMTs диметилируют белки, связанные с регуляцией транскрипции [2].

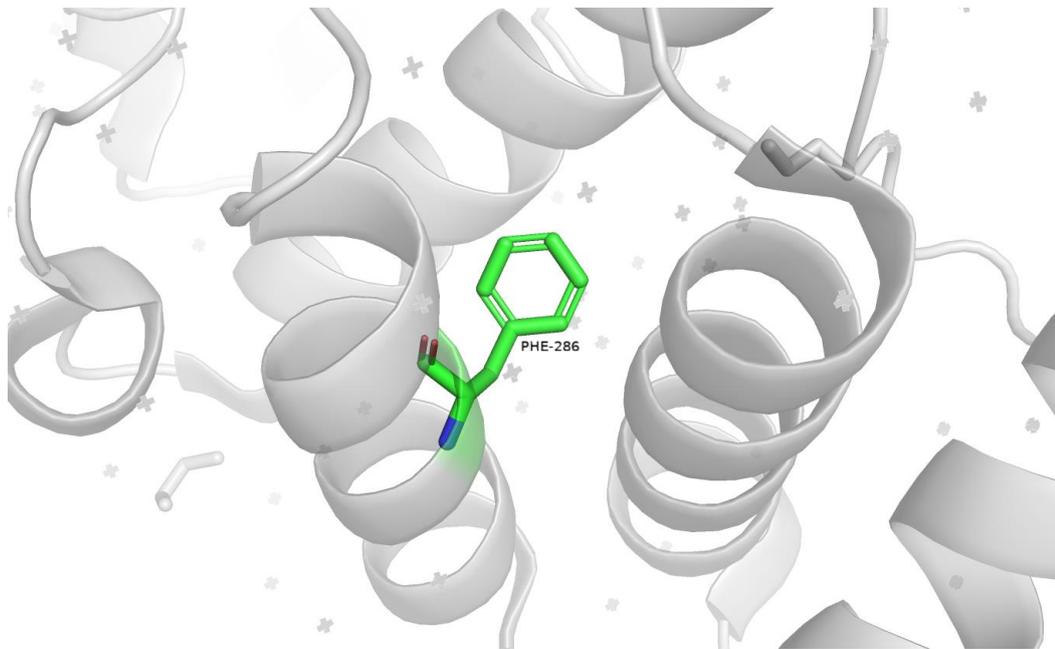


Рис. 2. Остаток Phe-286 (с наименьшим В-фактором)

Таким образом, остаток цистеина, находящийся на петле ближе к наружной части белка, имеет высокое среднее значение В-фактора, а остаток фенилаланина, расположенный на α -спирали в глобулярной части белка – низкое среднее значение В-фактора. На Рис. 3 показаны оба остатка - Cys-245 и Phe-286 – на фоне всего белка.

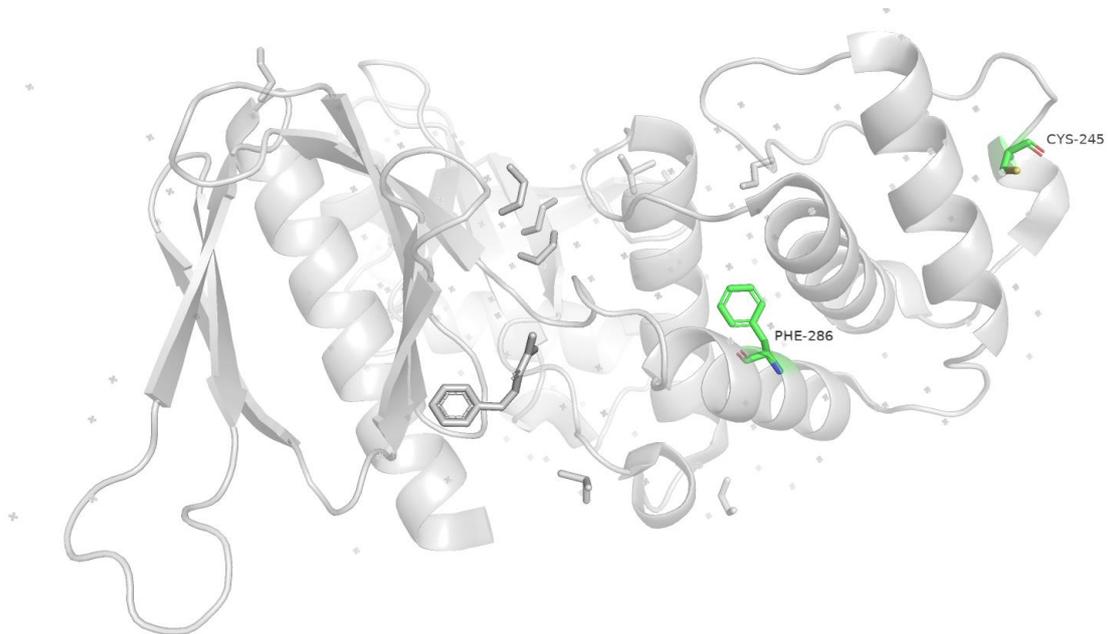


Рис. 3. Положение остатков Cys-245 и Phe-286

Задание 2. ProDy и В-факторы часть 2

С помощью ProDy был вычислен центр масс белка и посчитаны расстояния от него до центров масс каждого остатка. Для каждого остатка я получила зависимость расстояния от среднего значения В-фактора (см. Рис. 4).

На графике видно, что B-фактор остатка растет с увеличением его расстояния до центра масс белка. В общих чертах, это соотносится с утверждением о том, что внутри белковой глобулы остатки менее подвижны, чем на наружной части белка.

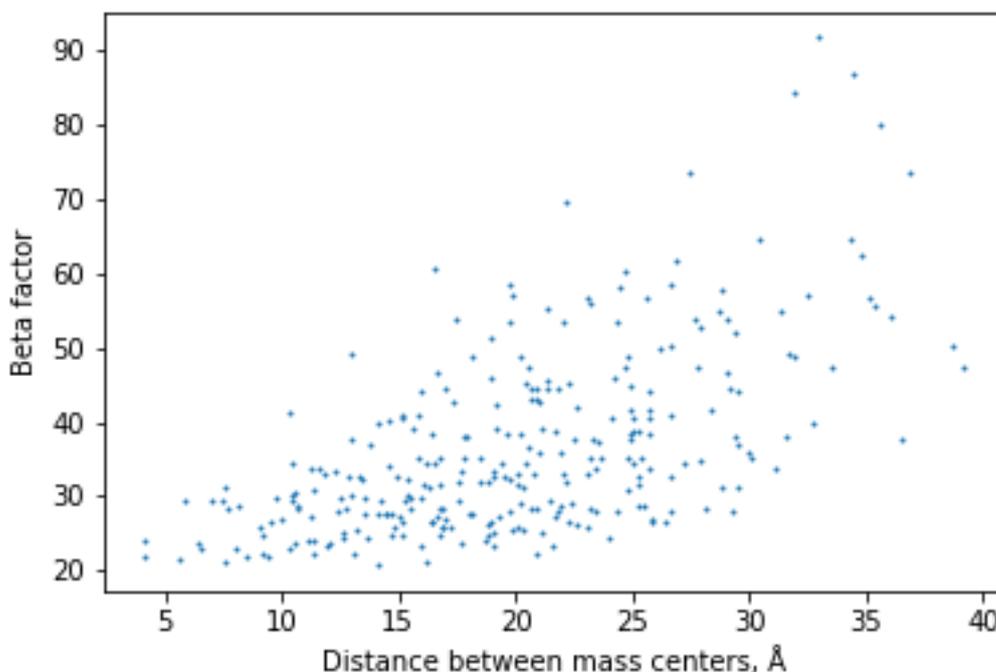


Рис. 4. Зависимость B-фактора остатка от расстояния до центра масс белка.

Задание 3. Как работает восстановление функции электронной плотности по экспериментальным данным

Была сгенерирована модельная функция ЭП в одномерной элементарной ячейке.

Вот какие параметры были выбраны:

```
python3 compile-func.py -g 2,3,2.8+25,3,4+2,3,5.1+2,3,8.9+20,3,10+30,3,15+25,3,16.5
```

Это сумма семи гауссовых функций: их максимумы в точках 2.8, 4, 5.1, 8.9, 10, 15, 16.5, высота 2, 25, 2, 2, 20, 30, 25 соответственно, ширина гауссиана 3. Выдуманная молекулярная система состоит из трех молекул: одна из них трехатомная, две – двухатомные. Далее был произведен расчет гармоник Фурье. Для "отличного" восстановления было достаточно первых 28 гармоник (см. Рис.5).

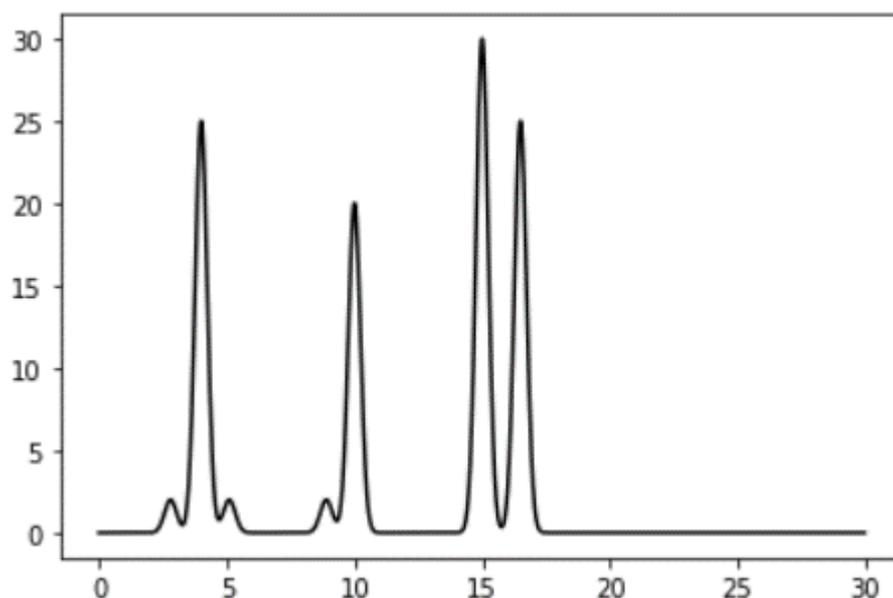


Рис. 5. «Истинная» электронная плотность выдуманной молекулярной системы.

Качество восстановления определялось следующим образом:

- Отличное восстановление – по графику восстановленной функции можно определить положение максимума всех гауссовых слагаемых функции ("атомов")
- Хорошее восстановление – можно угадать положение всех максимумов, зная число слагаемых ("атомов"), хотя на восстановленной функции максимумы от атомов не отличимы от шума
- Среднее восстановление – положение каких-то атомов определить по восстановленной функции нельзя, других - можно
- Плохое восстановление – положение атомов определить не представляется возможным; можно только предсказать примерный размер "молекулы"

Тетрадку с отчетом по заданию можно скачать по [ссылке](#). Ниже представлена таблица восстановления функции по коэффициентам Фурье.

Разрешение полного набора гармоник – это период гармоник с наибольшим номером. Оно определяется как длина отрезка, деленная на номер гармоник.

Таблица 1. Восстановление функции по коэффициентам ряда Фурье.

Набор гармоник	Разрешение (Å)	Полнота данных (%)	Шум амплитуды (% от величины F)	Шум фазы (% от величины phi)	Качество восстановления (отличное, хорошее, среднее, плохое)	Комментарии
Полный набор гармоник						
0–1	30 Å	100%	0	0	Ужасное	Даже размер молекулы определить невозможно
0–3	10 Å	100%	0	0	Очень плохое	Даже размер молекул определить невозможно, хотя уже понятно что молекул тут больше одной, раз пик раздвоился

0-10	3 Å	100%	0	0	Плохое	Можно только определить размер молекул
0-15	2 Å	100%	0	0	Среднее	Пики более тяжелых атомов различаются, однако положение атомов водорода по пикам определить нельзя
0-20	1.5 Å	100%	0	0	Среднее	Пики более тяжелых атомов уже близки к "истинным", однако положение атомов водорода нельзя определить из-за шума
0-25	1.2 Å	100%	0	0	Хорошее	Пики водородов уже можно угадать, но кое-где присутствует шум, сравнимый с сигналом водородов (у шума между пиками тяжелых атомов амплитуда больше, чем у сигнала "истинных" водородов. молекулярная система простая, поэтому положение водородов все же можно угадать, хоть и с трудом)
0-28	1.1 Å	100%	0	0	Отличное	Не идеально, но все сигналы хорошо различимы, шума почти нет
0-28	1.1 Å	100%	30	0	Среднее	Более тяжелые атомы можно распознать, водороды - нельзя
0-28	1.1 Å	100%	0	30	Среднее	Пики более тяжелых атомов угадываются, но с трудом, водородов не видно
0-28	1.1 Å	100%	15	15	Хорошее	Можно распознать пики водородов, если знать их количество
0-28	1.1 Å	100%	15	0	Почти отличное	Пики водородов оказались несколько завышены, что позволяет отличить их сигналы от шума
0-28	1.1 Å	100%	0	15	Среднее	Водороды не определить
Неполный набор гармоник						
2-28	1.2 Å	93%	0	0	Почти отличное	Сигналы уехали вниз, но все максимумы можно отличить, хоть и с трудом
0-15, 20-28	1.1 Å	82%	0	0	Среднее	Из-за шума водородов не увидеть, но пики более тяжелых атомов различить можно
0-15, 20-28, 35	1.1 Å	65%	0	0	Среднее	Ситуация практически идентична предыдущей
0-28, 35	1.1 Å	82%	0	0	Хорошее	Водороды можно распознать, но их пики почти на уровне шума
0-15, 20-28, 35-40	1.1 Å	70%	0	0	Среднее	Водороды не распознать

Источники

1. <https://www.uniprot.org/uniprot/Q9Y2J2> - о белке со структурой 5RYR
2. https://en.wikipedia.org/wiki/Protein_arginine_methyltransferase_5 - о PRMT5